Department of Statistics: PhD Presentation Event

Thursday 23rd and Friday 24th June 2011
Leverhulme Library (COL.B617), Columbia House

Abstracts

Daniel Bruynooghe	Differential cumulants, hierarchical models and monomial ideals

This talk is about local dependence measures, conditional independence statements and their non-parametric estimation, graphical and hierarchical models and algebraic statistics. The centre stage of the story is occupied by the differential cumulants. Whilst closely akin to their global counterparts, these specimens have never quite received the attention they deserve. Only when they are nowhere to be seen do they unfold their enormous powers. It is in those times that their isomorphic counterparts, the conditional independence statements, resurface. Fortunately, not all is lost, and it is the striking absence of the differential cumulants which makes them estimable using kernel methods. Time permitting, we cross the bridge to algebra and illustrate some fun connections.

Dan Chen	Stochastic volatility of volatility

This study considers two kinds of volatility of volatility: SPX's volatility of volatility and VIX's volatility. I proposed several SVV models, found out the methods to estimate them and compare those SVV models with the SV models.

Zhanyu Chen	Put-call symmetry in stochastic volatility models

The classic *put-call symmetry* has been introduced by Bowie and Carr in 1994 and successfully applied to the semi-static hedging of the barrier options. Afterwards, properties and conditions for the symmetric processes have been analysed by Carr and Lee, Farjado, Schmutz and Rheinlander in recent years. Nevertheless, the empirical results indicate that in general markets are not symmetric. In this study, we work on the extension of the *put-call symmetry* to the stochastic volatility model with correlation between the return of price process and the instantaneous volatility

Yehuda Dayan	TBC

TBC

Joseph Dureau	Capturing the time-varying drivers of an epidemic

The aim of this work has been to provide with an efficient inference methodology for time-varying parameters of non-linear equations, with a specific scope on epidemic modelling.

Time-variations of parameters are modelled by diffusion processes whose coefficients are unknown, in order to achieve both flexibility and parsimony. In the perspective of conservative public decisions, uncertainty on the diffusion coefficients and other biological parameters needs to be accounted for, suggesting a Bayesian approach commonly adopted in the field of epidemiology.

From a computational point of view, the problem is particularly challenging due to the high dimensionality and the very strong dependencies of the targeted posterior distribution. Other approaches have used approximations (Gaussian approximation, no noise approximation, etc…) which can provide with misleading results. We have implemented and adapted two recent methodologies, the Particle MCMC algorithm and the Maximum likelihood via Iterated Filtering that bring interesting solutions to the problem. Both these approaches will be presented and discussed.

This work opens the way to many applications both in theoretical epidemiology and public policies evaluation, an example of which will be quickly presented. Meanwhile, this methodology still remains computationally expensive and leaves space for further improvement.

Mai Hafez	Modelling dropout in longitudinal studies using latent variable models

The problem of missing data is very common in longitudinal studies where subjects often dropout prematurely before the end of a study. Ignoring the missingness mechanism can lead to biased estimates, especially when the missingness is "non-ignorable". This problem is treated using one of two approaches: selection models or pattern-mixture models. We present a literature review for the two approaches. Selection models are studied and previous results are replicated within this framework. Pattern-mixture models are also presented. Most of the literature on dropout in longitudinal studies deals with one variable over time. Our proposed approach uses latent variable models within the pattern-mixture models framework to account for dropout in multivariate longitudinal data. A latent variable is used to summarize an attitude or a construct at each time-point. Another latent variable is used to summarize tendency to dropout. Different relationships among these variables are proposed via three different model specifications.

Alex Jarman	Small-number statistics, common sense and profit: challenges and non-challenges for hurricane forecasting

When making only one forecast per year, or per decade, it can take some time to establish statistical confidence in the skill of a given forecast scheme. Must a risk tolerant decision maker wait decades until skill is "proven" if that decision maker believes the system to have value? A methodology is illustrated to demonstrate there are imperfect forecast systems which almost certainly have nontrivial value long before one might establish that their skill was statistically significant. The effects of serial correlation on forecast verification statistics are also empirically investigated.

Karolos Korkas	Adaptive estimation for piecewise stationary autoregressions

This presentation outlines a new method that relinquishes the assumption of stationarity in time series. Many time series in applied science are characterised by nonstationarity. Here, we deal with piecewise stationarity which is the simplest type of departure from stationarity. We develop a nonparametric method that estimates time-varying AR(1) coefficients of a nonstationary time series. For this purpose we adapt the Fused Lasso method of Tibshirani et al (2005) and we solve this high-dimensional penalised least squares problem by using a recently-developed path algorithm by Tibshirani and Taylor (2011). We present simulation studies which indicate that our method is capable of fitting AR(1) processes locally even when there are more than one breakpoints in the sample.

Jia Wei Lim	Distribution of the Parisian stopping time

We use a new method to invert the Laplace transform of the Parisian stopping time, which allows us to look at the asymptotic distribution of the stopping time. The results can be used to price Parisian options.

Malvina Marchese	Asymptotic distribution of some common panel estimator in large mixed panels

This paper studies the asymptotic properties of standard panel data estimators in a regression model with mixed stationary and nonstationary regressors, where the disturbances are allowed to be stationary or nonstationary. We show that all the estimators have asymptotic normal distribution but different rate of convergence depending on the nonstationarity of the regressors and the disturbances. In finite sample we show with a Monte Carlo experiment that GLS over perfoms in efficiency all other estimators.

Felix Ren	The methodology flowgraph model

Flowgraph models are one type of multistate stochastic process model that are used to describe time-to-event data, they give graphical representation of dynamic changes in stochastic system. Flowgraph model consists of nodes representing the states and directed line segments called branches indicating the direction of possible state changes. The objective of flowgraph analysis is to determine the distribution of the total waiting time between two nodes of interest.

The concepts of flowgraph theory and Mason's rule allow us to develop closed expressions for the moment generating function (MGF) of total waiting time between two nodes of interest. We propose a new derivation of the Mason's rule based on the inter-nodes transition matrix of flowgraph, and illustrate the method in a complicated flowgraph example. Padé approximation is introduced to approximation MGF in rational form, which can be applied to give closed form estimation for the probability density function of total waiting time. For parameter estimation, we compare the Maximum Likelihood estimator (MLE) with the Method of Moment (MOM). The bias correction method for MOM is also presented.

Roy Rosemarin	Projection Pursuit Conditional Density Estimation

Projection Pursuit approximations are known to offer a method to avoid the curse of dimensionality by a series of projections of high dimensional data into univariate directions. In this talk I consider a nonparametric conditional density projection pursuit approximation for a random variable y given a high-dimensional random vector x, where the approximation is chosen to minimise the Kullback-Leibler relative entropy between the true and the estimated density. We then suggest an iterative conditional density kernel-based estimation procedure for the projection pursuit approximation, and derive the asymptotic properties of the estimators

Ilya Sheynzon	Multiple equilibria and market crashes

To the best of our knowledge, all existing multiple equilibria and market crashes models are just one step models and explain only the nature of market crashes, rather than analyse the dynamics in terms of the distributions. We consider four alternative models (exogenous shocks; close to a barrier switching; stochastic number of dynamic hedgers; Bessel bridges approaches). Within this framework, we have theoretical models that allow us to analyse and predict the dynamics of the equilibrium price process in a continuous time. Then we discuss how we can use the market microstructure theory and Bayesian inference to estimate the parameter values and how we can deal with the discretely sampled data in case it is sampled relatively often. Finally, we apply numerical techniques (PDE; Volterra integral equations and approximation by piecewise linear boundaries approaches) to estimate the distribution formulas.

Ilaria Vannini	Multivariate regression chain graph models for clustered categorical data

Multivariate regression chain graph models are a flexible tool to analyse multivariate data. Making use of recent developments in marginal modelling, we propose a new methodology for fitting and testing such models when categorical clustered responses occur. The general parameterisation of log-linear marginal modelling is described and linked with maximum likelihood estimation procedure using an algorithm which overcomes some of the difficulties of other MP estimation methods. We present an illustration aimed at analysing couples' fertility intentions based on the 2003 Italian Family and Social Subjects Survey.

Ed Wheatcroft	Forecasting the meridionial overturning circulation

The meridionial overturning circulation (MOC) is circulation of water in the oceans. Differences in temperature and salinity cause different densities of water at different latitudes on Earth. Dense water sinks and is displaced by warmer water causing a circulation. This talk will discuss the well-known Stommel model and its different states. We briefly look at the MOC in other more complex models. We then go on to discuss the implications of chaos and use the Moran-Ricker map as an example. Finally we consider the method of shadowing and how it can help us in relating climate models and reality.

Yang Yan	Co value-at-risk measure

The Financial crisis of 2007 to 2009 has generated interest in systemic risk by people in industries, regulators and academics for both the theoretical and practical reasons. During the financial crises, losses tend to spread across financial institutions, and such increase of comovement between financial sectors give rise to systemic risk, which is threatening the financial system as a whole. Hence, it is important to develop formal measures of systemic risk, a good measure should serve as an early warning signal of potential market dislocation and capture the increase of tail comovement during financial distress. Our paper is most related to the systemic risk measure, CoVaR, proposed in Adrian and Brunnermeier (2010). By extending their work to Econometric methods, the risk measure can be first estimated by quantile regression and the standard errors can be corrected. A crucial extension is to build up the model into a simultaneous equation system, from which the dynamics of individual and market returns can be captured more accurately.

Hongbiao Zhao	Risk process with dynamic contagion claims

We consider a risk process with the intensity of claim arrivals modelled by a combination of Cox process and Hawkes process, i.e. the dynamic contagion process newly introduced by Dassios and Zhao (2011). This process is reviewed and some key problems in risk theory such as the generalized Lundberg's fundamental equation, net profit condition and ruin probability based on this new risk model are studied with the aid of martingale approach and change of measure.

Poster abstracts

Xiaonan Che	Stochastic boundary crossing probability for Brownian motions

We provide an alternative method to obtain the explicit formulae for the distribution of the first-passage time of a Brownian motion through linear boundaries, to which extent the unconditional probability for the Brownian motion to reach one of two stochastic boundaries is derived, particularly compound Poisson boundaries with exponential changes. Some powerful tools from the martingale theory are applied. The applications of the results, especially on payment system, are given. Both numerical examples and simulated resulted are provided.

Joseph Dureau	Inference on epidemic models with time varying parameters: methodology and preliminary applications

The dynamics of epidemics are quite frequently modelled by deterministic ordinary differential equations (ODE), the parameters of which reflect some behavioural or biological mechanisms. Such mechanisms depend on a number of factors, for example climate and behaviours, that vary across time, thus influencing the epidemic dynamics. Things are further complicated by the fact that the impact of these mechanisms can only be partially observed, quantified and understood. Our proposed modelling framework considers stochastic extensions of the traditional ODE models by allowing some of their parameters to be time continuous paths of stochastic processes. We develop a data-driven inference procedure to estimate the time variations of transmissibility utilising the Particle Monte Carlo Markov Chain algorithm. This methodology is illustrated on simulated data as well as on the Google FluTrend data for France during both a “classic” seasonal epidemic (2008-2009) and the H1N1 pandemic (2009-10).

Flavia Giammarino	Indifference pricing with uncertainty averse preferences

We consider the indifference valuation of an uncertain monetary payoff from the perspective of an uncertainty averse decision-maker. We study how the indifference valuation depends on the decision maker's comparative uncertainty attitudes, and we obtain a characterization of increasing, decreasing, and constant uncertainty aversion in terms of cash-subadditive, cash-superadditive, and cash-additive quasiconvex risk measures.

Sarah Higgins	Blending ensembles from multiple models

The most accurate seasonal weather forecasts combine multiple models developed by different countries using equal weights. A methodology to blend each model’s forecasts using weights determined by the skill of each model is examined. As there is only a small forecast-outcome archive available for seasonal forecasts we look at combining multiple imperfect models from a non-linear system using a proper skill score to determine the weights.